

REMARKS/ARGUMENTS

These remarks are submitted in response to the Office Action of February 21, 2007 (Office Action). As this response is timely filed within the 3-month shortened statutory period, no fee is believed due.

Claim Rejections – 35 USC §101

At page 2 of the Office Action, Claims 1-23 were rejected under 35 U.S.C. § 101, it being stated that the claims define non-statutory processes because they merely recite manipulation in the abstract without a claimed limitation to a practical application.

Applicants respectfully submit that, as the Interim Guidelines on Patentable Subject Matter states, a practical application is claimed if the claimed invention physically transforms an article or physical object to a different state or thing, or if the claimed invention otherwise produces a useful, concrete, and tangible result.

Applicants further respectfully submit that the claimed invention clearly produces a useful, concrete, and tangible result, namely, that of generating a concatenative text-to-speech voice built using a set of verified phonetic units. The concatenative text-to-speech voice produced according to the present invention can be used in a variety of speech recognition devices to improve the quality of the speech, which is a tangible result.

It is also stated in the Office Action, that it is not clear what type of data is being received and how that data is received. Applicants respectfully disagree and expressly point out that the claims clearly recite the type of the data received: "*at least one phonetic unit.*" A person skilled in the art would readily know how such data would normally be received into a system. For example, phonetic units are automatically extracted from a speech corpus and received into a filtering system as inputs.

Although it is believed that the each of the claims thus clearly defines a patentable invention, the language of the claims has been amended to avoid the rejections. The amendments are fully supported in the Specification. No new matter has been introduced

through the claim amendments. Applicants, therefore, respectfully request that the rejections under 35 U.S.C. § 101 be withdrawn.

Claim Objection

At page 3 of the Office Action, Claim 4 was objected to because, as stated, it is not clear as to what the Applicant means by "different suspect phonetic unit." Applicants respectfully disagree and note that Claim 4 depends from Claim 3, which in turn, depends from Claim 2. Claim 2 expressly recites that a phonetic unit is marked as a suspect phonetic unit if the abnormality index exceeds the normality threshold. Accordingly, a suspect phonetic unit is one that has been marked as such because the calculated abnormality index of the phonetic unit exceeds a normality threshold. All the suspect phonetic units are stored in the suspect data store (140). The navigation control in the alignment validation interface (150) navigates from one suspect phonetic unit to another suspect phonetic unit ("*a different suspect phonetic unit*") in order for the alignment validation interface (150) to validate or reject the suspect phonetic unit. Applicant thus respectfully requests that the claim objection be withdrawn.

Claim Rejections – 35 USC §102

At pages 3-10 of the Office Action, Claims 1-23 were rejected. Each of the claims was rejected under 35 U.S.C. § 102(b) as being anticipated by U.S. Patent No. 6,665,641 to Coorman, *et al.* (hereinafter Coorman).

Aspects Of Applicant's Invention

It may be useful to reiterate certain aspects of Applicants' invention prior to addressing the cited reference. The present invention is directed to a method and a system for detecting misaligned phonetic units from phonetic units to be used within a concatenative text-to-speech voice.

One embodiment of the invention, typified by Claim 1, is a method of filtering phonetic units to be used within a concatenative text-to-speech voice. The method can include receiving into a filtering system at least one phonetic unit that has been automatically extracted from a speech corpus in order to construct a concatenative text-to-speech voice. The method also can include calculating an abnormality index for the phonetic unit, wherein the abnormality index indicates a likelihood of the phonetic unit being misaligned, and comparing the abnormality index to a normality threshold. According to the method, if the abnormality index does not exceed the normality threshold, the phonetic unit can be marked as a verified phonetic unit. The method further can include building a concatenative text-to-speech voice using the verified phonetic units.

The Claims Define Over the Prior Art

Coorman is directed to a corpus-based speech synthesizer in which the speech is generated from a large database of continuous speech, which has not been segmented and filtered to exclude misaligned phonetic units. With Coorman, the selection of the phonetic units is performed at the run time. With Applicants' invention, by contrast, the concatenative text-to-speech voice is pre-selected and consists only of verified phonetic units, all misaligned phonetic units having been excluded. The various advantages of the present invention thus include reducing the size of the voice database and improving the quality of the speech.

Applicants respectfully note that Tables 6 and 7 of Coorman, cited in the Office Action, are examples of "cost functions." (See, Col. 13, lines 33-38.) Coorman explicitly uses the cost functions to select waveforms from the speech database. (See, Col. 13, line 38 Col. 16, line 22.)

Applicants respectfully note that, as described, Coorman's cost functions are used to determine how well candidate waveforms can be joined together. But this is wholly

unrelated to, and provides no mechanism for, excluding misaligned phonetic units.

Corman's cost functions are in no way comparable to an abnormality index. It follows, therefore, that Coorman does not calculate an abnormality index for phonetic units. Accordingly, Coorman fails to compare a computed abnormality index of a phonetic unit to a normality threshold in order to determine whether the phonetic unit is to be marked as a verified phonetic unit. It thus further follows that Coorman is incapable of building a concatenative text-to-speech voice using only verified phonetic unit; that is, one that excludes misaligned phonetic units from the speech database to be formed, as with Applicants' invention.

Coorman thus does not expressly or inherently teach every feature recited in the claims. Applicants, therefore, respectfully submit that Claims 1-23 define over the prior art.

CONCLUSION

Applicants believe that this application is now in full condition for allowance, which action is respectfully requested. Applicants request that the Examiner call the undersigned if clarification is needed on any matter within this Amendment, or if the Examiner believes a telephone interview would expedite the prosecution of the subject application to completion.

Respectfully submitted

Date: May 21, 2007



Gregory A. Nelson, Registration No. 30,577
Richard A. Hinson, Registration No. 47,652
Yonghong Chen, Registration No. 56,150
AKERMAN SENTERFITT
Customer No. 40987
Post Office Box 3188
West Palm Beach, FL 33402-3188
Telephone: (561) 653-5000